International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/


(RESEARCH ARTICLE)



# Forecasting currency exchange rates using EMD-ARIMA Model

Dennis Cheruiyot Kiplangat *

*Department of Statistics and Actuarial Sciences, Dedan Kimathi University of Technology, Kenya.*

International Journal of Science and Research Archive, 2024, 13(02), 1469–1484

Publication history: Received on 13 October 2024; revised on 24 November 2024; accepted on 26 November 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.2.2273

## Abstract

In today's global economy, accuracy in forecasting currency exchange rates is of much importance to any future investment. Currency exchange rates portray non-linear and non-stationary characteristics hence to address these characteristics; this paper proposes a hybrid forecasting model using the Empirical Mode Decomposition (EMD) technique, and the ARIMA model. EMD is used to decompose the raw currency exchange rate data into several intrinsic mode functions and one residual. The process of extracting the IMFs from the data is called the sifting process. EMD was used to detect the moving trend of currency exchange rate data and improve the forecasting success of the ARIMA model. The data were obtained from the Central Bank of Kenya website between the periods January 2005 to May 2017. The best ARIMA model fitted to the raw data before decomposition based on information criterion statistics was found to be ARIMA (1,0,3) for the KShs/AE. Dirham, ARIMA (1,0,1) for KShs/Australian dollar and ARIMA (1,0,3) for KShs/Canadian dollar currency exchange rates. After forecasting, we then compared the forecasted values with the actual data to check the suitability of the ARIMA model. Further, EMD was applied to the exchange rate data and then fitted an ARIMA model to the IMFs. The best model was found to be ARIMA (1,0,1) for the KShs/AE. Dirham, ARIMA (0,0,1) for KShs/Australian dollar and ARIMA (1,0,1) for KShs/Canadian dollar currency exchange rates. The appropriateness of these models was tested using the Ljung-Box test. The forecasting performance of each model was evaluated using the RMSE.

**Keywords:** Autoregressive Integrated Moving Average (ARIMA); Empirical Mode Decomposition (EMD); Intrinsic Mode Function (IMF), Quantile-Quantile plot, Root Mean Squared Error (RMSE), Currency Exchange Rates.

## 1. Introduction

An exchange rate refers to the currency rate of a particular country expressed in terms of the currency of another country. Currency exchange rates are one area of interest for financiers and investors. How to forecast currency exchange rates in the future is one subject that all financiers and investors care about when planning. The more accurate the forecast result is, the more effective the plan is. The constant movement of currency exchange rates combined with the rapid globalization of business has resulted in the demand for forecasting methods.

Financial time series forecasting plays an important role in the world's economy due to its ability to forecast economic benefits and influence countries' economic development. International currency trading is a crucial economic index for international trade, financial markets, the alignment of economic policy by governments and corporate financial decision-making. However, it is known that financial time series forecasting has shortcomings such as non-linearity and non-stationarity. Therefore, financial time series forecasting is a challenging task in financial markets.

In this study, the Autoregressive Integrated Moving Average model is used to forecast currency exchange rates. When modelling financial time series using ARIMA models we have to consider that financial time series data is non-linear, noisy and non-stationary. Therefore, we first transform the raw data using Empirical Mode Decomposition and then

* Corresponding author: Kiplangat Dennis Cheruiyot

forecast currency exchange rates using the ARIMA model. EMD is suitable for financial time series data in terms of finding fluctuation tendency, which then simplifies the forecasting task into several forecasting sub-tasks. EMD as a data processing technique offers a way by which the non-stationary and non-linear characteristics of raw currency exchange rate data can be decomposed into several Intrinsic Mode Functions. It also reveals the hidden patterns and trends of time series.

ARIMA model coupled with EMD may present a financial time series forecasting model for currency exchange rates, where consideration of the decomposed financial time series structure is expected to increase the accuracy of the proposed model in terms of overcoming the non-linearity and non-stationarity limitations. The following exchange rates were used in the study; KShs/AE. Dirham exchange rate, KShs/Australian Dollar exchange rate and KShs/Canadian Dollar.

### 1.1. Problem Statement

The Exchange rate prediction is a challenging application of modern time series forecasting. The currency exchange rates are noisy, non-stationary and non-linear. These characteristics suggest that there is no complete information that could be obtained from the past behavior of currency exchange markets to fully capture the dependency between the future rates and that of the past. A model needs to be fitted to accurately forecast the future currency exchange rates. In the recent past, researchers have attempted to forecast currency exchange rates using various models like ARIMA model. In the quest to improve the prediction accuracy, this paper proposed a method of decomposing the currency exchange rate data into several Intrinsic Mode Functions which are then forecasted using the ARIMA model.

### 1.2. Justification of the Study

Previous studies suggest that variations in exchange rates have the potential to affect a country's economic performance. Less developed countries, for instance, Kenya, have received less attention compared to those with developed economies. An appropriate exchange rate has been an important factor for the economic growth in the economies of most developed countries while an appropriate exchange rate has been a major obstacle to the economic growth of many African countries, Kenya being one of them.

### 1.3. Significance of the Study

An accurate forecast of currency exchange rates is important as the exchange rate affects the financial market and the development of the economy. Appropriate forecasting of currency exchange rates is of much importance for the success of many businesses and fund managers. For firms trading with other countries, movements in the currency exchange rates are of great importance. Firms may need to plan by hedging currency exchange rate movements or seeking to sell to domestic markets. For instance, appreciation in Kenyan currency could reduce demand for exports.

Risk managers use currency exchange rate forecasts to measure potential risks that can arise from the market and devise various ways of managing those risks. Forecasting of currency exchange rates is vital in risk management since risk management is all about measuring potential losses of a portfolio and being able to estimate potential losses. Investors can also use currency exchange rate forecasts to measure the amount of risk associated with a certain combination of investments in a portfolio.

The currency forecast is an important element of financial markets and managerial decision-making. It is well known that using a weak forecasting technique in the financial market is harmful to economic development. Changes in currency exchange rates have an impact on international trade and foreign investments. Hence, improved forecasting accuracy is a matter of great importance.

## 2. Literature review

Forecasting of currency exchange rates is an extremely important concept in finance for numerous reasons. The literature on currency exchange rates agrees on the key phenomenon. There is evidence of movements in currency exchange rates, where all participants in the foreign exchange rates market such as investors, traders, fund managers and risk managers have consensus about it. This has led to researchers, authors and academics developing various models to forecast currency exchange rates.

Mong et al., (2016) used the ARIMA model in forecasting the Vietnam Dong versus US dollar (VND/USD) exchange rate. They introduced the ARIMA model with four steps to forecast foreign exchange between VND/USD in the next twelve months of 2016. In their study, they used real foreign exchange data from the year 2013 to year 2015. After forecasting, they then compared them with real foreign exchange data to check the suitability of the ARIMA model in forecasting currency exchange rates. The results of the study showed that ARIMA is suitable for short-term forecasting of currency exchange rates.

Nwankwo, (2014) used the ARIMA model for analysis of the currency exchange rate between the Nigerian Naira and US dollar from the year 1982 to the year 2011. He carried out the analysis to help those who are interested in the irregular trend of the Nigerian exchange rate system. Different models were fitted though AR (1) and AR (2) had the same structural features, the best fit was the AR(1) model because it had the most suitable AIC. This was achieved through the diagnostic checking which identified it as the best fit.

ARIMA model was used by Appiah et al., (2011) in modelling the monthly exchange rate between the Ghana Cedi and US Dollar and forecasting the future currency rates. They forecasted the exchange rate to see if it could be appreciated to bring out improvement in trade imbalance to better Ghana's economy. The results of their study showed that the predicted rates were consistent with the depreciating trend of the observed series. Thus, ARIMA model was found to be the most suitable model.

Empirical Mode Decomposition was used by Hong, (2011) in the decomposition and forecasting of financial time series with high frequency. In his study, the EMD of the wavelet transformation was introduced into the processing of financial time series with high frequency, where the high-frequency data was first decomposed with EMD. The development trend of each component of IMFs was then explored in different time scales. The forecast model was finally reconstructed by using IMF components, which were used to analyze the time series of oil futures at 5-minute intervals. The results of his study showed that a more accurate forecast can be made with EMD by extracting the IMF components with different trends in the financial time series.

Zeng et al., (2014) employed an approach for Baltic dry index analysis based on EMD. The bulk shipping market has the characteristics of seasonality and cyclicality. Due to the non-stationary and nonlinear price series and the complexity of influencing factors, it is difficult to analyze the fluctuation of the bulk shipping market. In their study, a method based on empirical mode decomposition (EMD) was proposed to investigate the characteristics of the Baltic Dry Index (BDI). In the method, the original freight price series was decomposed into several independent intrinsic modes using EMD first. Then the intrinsic modes were composed into three components (time series), namely, short-term fluctuations caused by normal market activities, the effect of extreme events, and a long-term trend. Numerical experiments indicated that the proposed method effectively revealed the characteristics of bulk freight price series with different economic meanings, and decreased the error accumulation. Meanwhile, by the composition of intrinsic modes, the complexity of model formulation was controlled and the operation of the model was improved.

Motivated by the success of combining a particular model with a certain technique in forecasting financial time series data, in our study we will combine the ARIMA model with EMD. EMD which was founded by Huang et al (1998), is a technique that decomposes non-linear and non-stationary data by using the Hilbert –Huang Transform (HHT). It is suitable for financial time series to find the tendency of fluctuation, therefore simplifying the forecasting task into a few simple forecasting subtasks. EMD assist in designing forecasting models for various applications as it can reveal the hidden patterns and trends of financial time series.

Many previous researchers have proposed various hybrid models to ensure accuracy in forecasting financial time series. For example, Lin et al., (2012) used hybrid ARIMA and SVM models in forecasting stock prices. Yang and Lin combined the Empirical Mode Decomposition and neural networks model is better compared to a single model in forecasting financial time series. A hybrid model is developed to overcome the limitations of the single models and ensure more accuracy in forecasting financial time series.

In the case study of exchange rate forecasting using EMD and Least Squares Support Vector Machine, Izzati, Rashid, Samsudin, Shabri, and Komputeran, (2016), combined the LSSVM model with EMD to forecast daily USD/TWD exchange rate. The exchange rate data was first decomposed by EMD into several Intrinsic Mode Functions (IMF) and a residual. After that, LSSVM was used to forecast each of the groups and the forecasted values were summed up to obtain the final exchange rate forecasting value. Their result showed that the hybrid model of EMD-LSSVM outperforms the single LSSVM model.

Forecasting crude oil prices with an EMD-based neural network ensemble learning paradigm was a study done by Yu et al., (2008). They considered that crude oil price series are non-linear and non-stationary hence it was rather challenging to accurately predict the crude oil prices. They therefore used EMD to decompose the original crude oil price series into some independent IMF components. Each component was modelled by the neural network model such that the tendencies of the IMF components were accurately predicted. Prediction results of all IMFs were aggregated using the neural network to produce an ensemble forecasting result for the original crude oil price series. According to their results, they concluded that the neural network model combined with EMD is superior to a single neural network in terms of accuracy level of prediction.

In the recent past, forecasting of a financial time series has become an important area and has gained a lot of attention from researchers, investors and other interested parties hence the need to fit a model that accurately forecasts financial time series. The problem of forecasting currency exchange rates accurately has led to much research and the development of models which try to best forecast financial time series data.

## 3. Methods

Financial time series data which is non-stationary will be used in our study. A non-stationary data is where the mean and variance vary with time. The non-stationary data will be transformed to become stationary using the Empirical Mode Decomposition where the raw data will be decomposed into several components. Also, to make a financial time series stationary the differences between consecutive observations can be computed. Differencing can help stabilize the mean of a time series by removing changes in the level of a time series and so eliminating trends. On the other hand, transformations such as logarithms can help to stabilize the variance of a time series.

### 3.1. Empirical Mode Decomposition

Empirical Mode Decomposition is a non-linear, non-stationary data processing technique that was proposed by Huang et al (1998). It decomposes financial time series into several Intrinsic Mode Functions (IMFs) each having intrinsic time scale properties. To break up the original data into a series of IMFs, the EMD technique separates the data into a slow-varying local mean part and a fast-varying symmetric oscillation part. The oscillation part becomes the IMF and the local mean is the residue, the residue serves as the input data again for further decomposition. The process repeats until no more oscillation can be separated from the last residue. According to Huang et al, each IMF must satisfy the following two conditions; First, the number of extreme values and zero crossings are either equal or differ at most by one. Secondly, the mean value of the envelope constructed by the local maxima and minima is zero at any point. On each step of the decomposition, because the upper and lower envelopes are initially unknown, a repetitive sifting process is applied to approximate the envelopes and obtain the IMFs and residue as follows;

- Identify all the local minima and maxima of the time series data $M(t)$.
- Get the lower envelope $M_1(t)$. and upper envelope $M_\mu(t)$ of the $M(t)$.
- Calculate the first mean value $\mu_1(t)$, that is $\mu_1(t) = (M_1(t) + M_\mu(t))/2$.
- Evaluate the difference between the original time series $M(t)$ and the mean of the time series $\mu_1(t)$. The first IMF $q_1(t)$ is defined as $q_1(t) = M(t) - \mu_1(t)$.
- Check whether $q_1(t)$ satisfies the two conditions of an IMF property. If the two conditions are not satisfied, repeat the first three steps to find the first IMF.
- Once the first IMF is obtained, the above steps are repeated to find the second IMF until the final time series $\varepsilon(t)$ which is a residual component is obtained and fulfils the termination criteria which suggests stopping the decomposition procedure.

The original time series $M(t)$ can be obtained by summing all IMF components including the one residual component as follows;

$$M(t) = \sum_{i=1}^{n} q_i(t) + \varepsilon_m(t)$$

where;

$n$ is the number of IMFs.

$\varepsilon_m(t))$ is the final residue.

$q_i(t)$ are the IMFs which are nearly independent of each other and all have nearly zero means.

Thus, one can achieve the decomposition of the financial time series data into several empirical mode functions and one residue. The above sifting procedure can be implemented using R software. EMD can therefore be used as an effective decomposition method in the sense that it is relatively easy to understand and implement. Fluctuations within a time series are automatically and adaptively selected from the time series. It is also suitable for non-linear and non-stationary time series decomposition.

### 3.2. Moving Average

Moving Average is used with time series data to smooth out short-term fluctuations and highlight longer-term trends or cycles. It is a technique to get the overall idea of the trends in a dataset. The process $X_t$ is said to be a Moving Average process of order $q$ if;

$$X_t = \beta_0 Z_t - \beta_1 Z_{t-1} - \beta_q Z_{t-q}$$

where $\beta_i$ are constants. The random variables $Z_i's$ are independent and are usually scaled so that $\beta_0 = 1$. The Expectation of $MA(q)$ is given by;

$$E(X_t) = \sum_{j=0}^{q} \beta_j E(Z_{t-j}) = 0$$

The variance is given by;

$$Var(X_t) = \sum_{j=0}^{q} \beta_j^2 \, Var(Z_{t-j}) Var(X_t) = \sigma_Z^2 \sum_{i=0}^{q} \beta_i^2$$

### 3.3. Autoregressive (AR) Model

Autoregressive model is used with time series data where the future values are estimated based on a weighted sum of past values. When the value of a series at a current time period is a function of its previous values plus some error, the generated mechanism is an autoregressive process. The general equation of an $AR(p)$ process is;

$$X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \varepsilon_t$$

where; $X_t$ is the stationary dependent variable being forecasted at time $t$. $X_{t-1}, \dots, X_{t-p}$ is the response variable at time $t-1, \dots, t-p$ respectively. $\varepsilon_t$ is the error term at time t with mean zero and a constant variance, that is $\varepsilon_t \sim WN(0, \sigma^2)$. Using the backshift operator, the $AR(p)$ model is written as;

$$\left(1 - \alpha_1 B - \alpha_2 B^2 - \alpha_p B^p\right) X_t = \alpha(B) X_t = \varepsilon_t$$

### 3.4. Autoregressive Integrated Moving Average

ARIMA model is a time series model which applies autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) models to find the best fit of a time series model to past values of a time series. ARIMA model describes the current behavior of a time series with respect to its past values. ARIMA model can be decomposed into two parts (Box et al,1994). First, it has an integrated (I) component (d), which represents the amount of differencing to be performed on the series to make it stationary making it easier to forecast. The second part consists of an ARMA model for the series rendered stationary through differentiation. The model takes into account historical data and decomposes it into AR process, where there is a memory of past events and a MA of the forecast errors, such that the longer the historical data, the more accurate the forecasts will be, as it learns over time. The AR component captures the correlation between the current value of the time series and some of its past values. The MA component represents the weighted moving average over past errors. The model has the form;

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q}$$

where $Y_t$ and $\varepsilon_t$ are the actual value and random error at time t respectively. $\alpha_i$ and $\beta_i$ are the model parameters. $p$ and $q$ are integers greater than or equal to zero and are often referred to as the order of the model. Random errors, $\varepsilon_t$ are assumed to be independently and identically distributed with a mean of zero and a constant variance of $\sigma^2$.

### 3.5. Fitting the model

The model identification step involves identifying whether data is stationary or not. A time series is said to be stationary if both the mean and the variance are constant over time. Stationarity is necessary in building an ARIMA model that is useful for forecasting. A time plot of the data is done to determine whether any differencing is needed before performing formal tests. One of the ways of identifying non-stationary time series is the ACF plot. For a stationary time series, the ACF will drop to zero relatively quickly while the ACF of non-stationary data decreases slowly. If the data is non-stationary, we do a logarithm transformation or take the first (or higher) order difference of the data series which leads to a stationary time series. This process will be repeated until the data is rendered stationary. The number of times to difference the data is indicated by the parameter $d$ in the $ARIMA(p, d, q)$ model. The Augmented Dickey-Fuller Test (ADF Test) is used to determine the stationarity of the data.

It also involves identifying seasonality in the data and using plots of autocorrelation and partial autocorrelation functions of the data to determine which autoregressive or moving average components should be used in the model. The ACF represents the degree of persistence over respective lags of variables; and the correlation between two values of the same variable at time $X_t$ and $X_{t+k}$. PACF measures the amount of correlation between two variables. ACF will be used to identify the order of the MA process while PACF will identify the order of the AR model. For the AR process, the ACF tails off towards zero while the PACF cuts off to zero (after lag p). For the MA process, the ACF cuts off to zero (after lag q) while the PACF tails off towards zero. Therefore, ACF and PACF are necessary in providing ways to identify the order that the ARIMA model will take. There are three rules to be followed to identify the model which are; If the ACF graph cuts off after lag $n$ and PACF dies down we identify $MA(q)$ resulting in $ARIMA(0, d, n)$ model; If ACF dies down and PACF cut off after lag n, we identify AR(p) resulting in $ARIMA(n, d, 0)$ model; If ACF and PACF die down that is mixed ARIMA model, then differencing of the time series is needed.

Once the order of the ARIMA model has been identified, the parameters $\alpha_1, \alpha_2, \ldots.\alpha_p$ and $\beta_1, \beta_2, \ldots.\beta_q$ of the model are then estimated using the maximum likelihood method. The maximum likelihood estimation (MLE) will be used to estimate the parameters of the ARIMA model. The Maximum Likelihood Estimates are the values of the parameters that is most likely to have generated the observed sample of data. This method therefore finds the values of the parameters which maximize the probability of obtaining the data that we have observed. For ARIMA models, MLE is very similar to the least square estimates. Information criteria are likelihood-based measures of model fit that include a penalty for the number of parameters to be used. Different information criteria are distinguished by the form of the penalty and can prefer different models. The two commonly used information criteria are AIC and BIC. When fitting models, it is possible to increase the likelihood by adding parameters which may result in overfitting, where the model is excessively complex such as having too many parameters relative to the number of observations. Both AIC and BIC attempt to resolve this problem by introducing a penalty term for the number of parameters in the model. The penalty is larger in BIC than in AIC.

Under model checking, it involves checking whether the estimated model is appropriate, that is if the residuals are white noise. The residuals are required to be uncorrelated and normally distributed. If the estimated model is good, the residuals should satisfy these assumptions. If these assumptions are not satisfied, one needs to fit a more appropriate model by going back to the model identification step and trying to develop a better model. Model checking will be done using the Ljung-Box test or plotting autocorrelation and partial autocorrelation of the residuals.

### 3.6. Forecasting

Forecasting performance of the various types of ARIMA models will be compared by computing statistics like AIC, and Root Mean Square Error (RMSE). The smaller the statistics, the better the model. Lastly, we perform a diagnostic check to ensure the chosen model best fits. The Ljung-Box statistic indicates whether the model is correctly specified. Plotting the residuals of the estimated model is a useful diagnostic check by checking the white noise requirement of residuals.

### 3.7. Checking the forecast accuracy

RMSE is used as a measure to compare forecasts for the same series across different models. The smaller the error, the better the forecasting ability of that particular model according to the RMSE criterion. It is represented in the equation below where $\hat{Y}_i$ and $Y$ is the forecast value and actual value respectively and $n$ is the total number of data points.

## 4. Results

### 4.1. Descriptive Statistics

The dataset consists of three foreign exchange rates, that is, KShs/AE. Dirham, KShs/Australian Dollar and KShs/Canadian Dollar from January 2005 to May 2017 daily observations excluding weekends and public holidays. A financial time plot of the daily foreign exchange rates of the three currencies against the Kenya shilling is shown in Figure 1. It shows that the data is non-stationary which is confirmed by the Augmented Dickey Fuller test. The Augmented Dickey-Fuller test statistic for KShs/AE. Dirham is -2.8998 with a p-value 0.1974 which is greater than the conventional level of significance, 5% which implies that the data is non-stationary. For the KShs/Australian dollar the test statistic is -1.5892 with a p-value of 0.7522 which is greater than 5% and -2.033 and a p-value of 0.5644 for the KShs/Canadian dollar. These test results confirm that the three series of currency rates are non-stationary and require differencing to make them stationary.

This is done by applying natural logarithm to the returns of the currency rates. Thereafter the log returns are tested for stationarity and the results are positive as follows; for KShs/AE. Dirham, the test statistic is -11.736 with a p-value of 0.01 which is less than 0.05, the level of significance, for KShs/Australian dollar, the test is -15.102 with a p-value of 0.01 which is less than the 5% level of significance and for KShs/Canadian dollar the test statistic is -15.48 with a p-value of 0.01. The results imply that the null hypothesis of non-stationarity is rejected at a 5% level of significance.

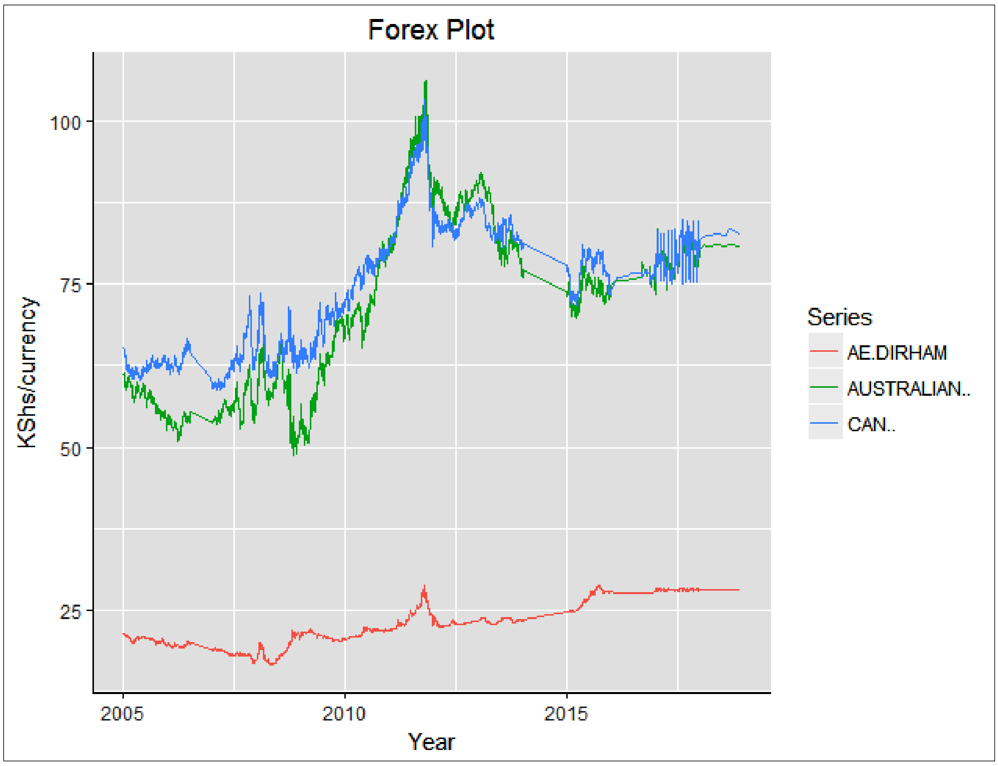


**Figure 1** Daily foreign exchange rates of Kenyan shillings versus the other three currencies

Table 1 presents the summary statistics of the three currency log return series. The number of observations as shown is 2,652. The mean for the three currencies is low, showing the varying disparity of returns. The variance of log returns is also low as the differences in currency exchange rates are minimal and time-varying. The kurtosis for the three series is greater than three implying that the log-returns of the currency rates are non-normal as is shown on the Quantile-Quantile (Q-Q) plots in Figure 2-4.

**Table 1** Summary statistics for the KShs/currency

| Currency | AE. Dirham | Australian Dollar | Canadian Dollar |
|---|---|---|---|
| Number of Observations | 2652 | 2652 | 2652 |
| Minimum | -0.056297 | -0.07473 | -0.078319 |
| Maximum | 0.04842 | 0.076513 | 0.098509 |
| Mean | 0.000087 | 0.000041 | 0.000035 |
| Median | 0.007600 | 0.000008 | 0.000087 |
| Variance | 0.000027 | 0.000130 | 0.000108 |
| Standard Deviation | 0.005222 | 0.011408 | 0.010394 |
| Skewness | -0.340693 | 0.106068 | 0.636221 |
| Kurtosis | 22.821714 | 6.91613 | 13.539774 |

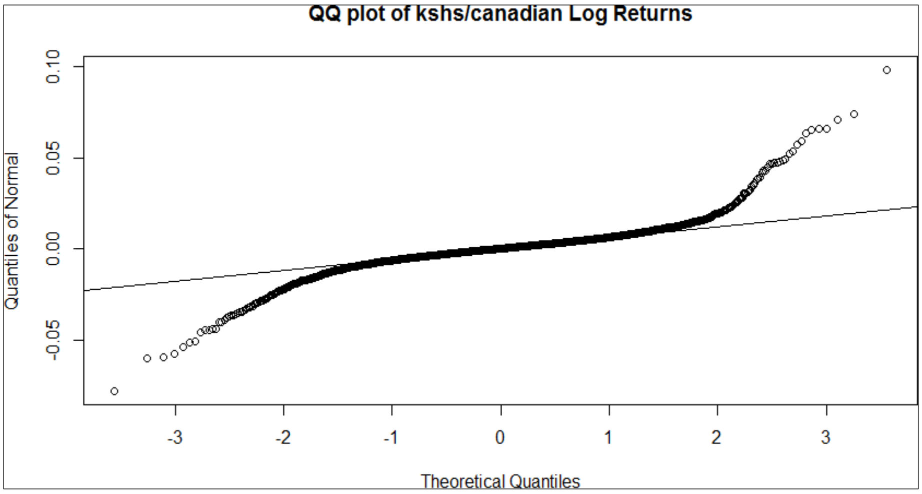


**Figure 2** QQ plot of KShs/Dirham Log-returns

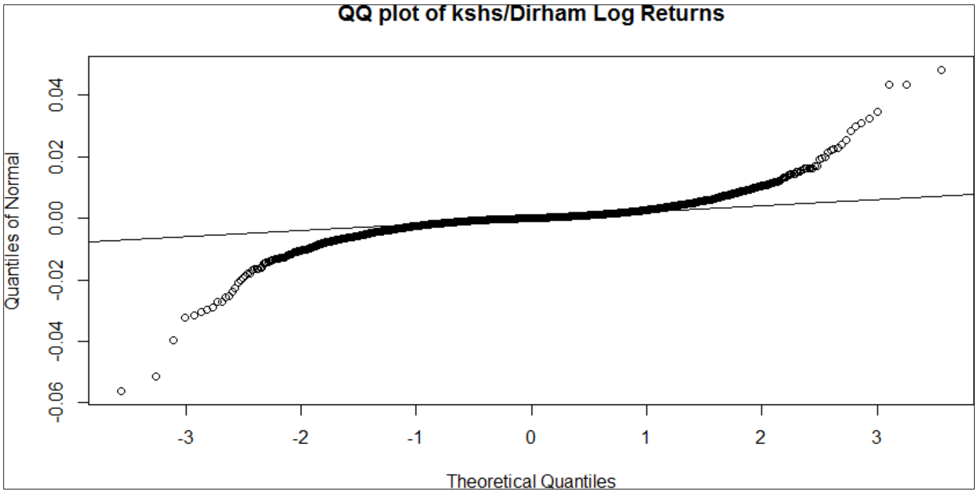


**Figure 3** QQ plot of KShs/Canadian dollar Log-returns

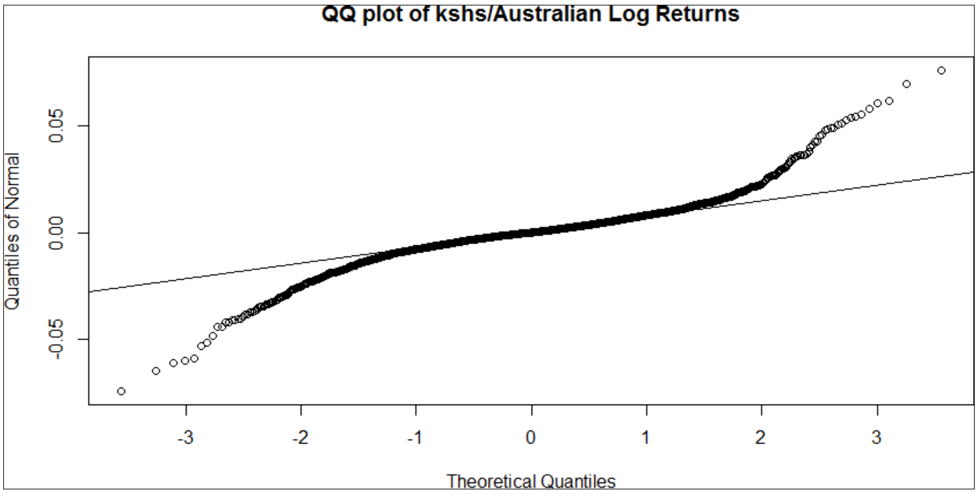


**Figure 4** QQ plot of KShs/Australian Log-returns

The Q-Q plots display a relationship between the normal quantiles and the underlying distribution's quantiles. From figures 2-4, it is evident that the log-returns of the currency rates are not normal. Quantitatively, the Jarque-bera test is applied to test for normality. The test result for KShs/AE. Dirham is 57701 with a p-value of 0.0000, for the KShs/Australian, the test statistic is 5301.9 with a p-value of 0.0000 and for KShs/Canadian dollar, the test statistic is 20474 with a p-value of 0.0000. This result leads to the rejection of the null hypothesis of normality. Further Auto Correlation Function (ACF) and Partial Autocorrelation Function (PACF) plots applied showed the presence of mean effects or serial correlation of the currency returns. The spikes of the ACF and PACF plots are sparsely distributed which shows the presence of weak serial correlation.

Alternatively, quantitative results are obtained from the Ljung-Box test which is applied to logarithmic returns of the currency rates to check for serial correlation. The results obtained are as follows; for KShs/AE. Dirham the test statistic is 77.94 with a p-value of 0.0000, for the KShs/Australian dollar, the test statistic is 52.68 with a p-value of 0.0000 and finally for the KShs/Canadian dollar the test statistic is 71.136 with a p-value of 0.0000. The results lead to the rejection of the null hypothesis of no serial correlation since the test statistics are large with p-values less than the level of significance of 5%.

### 4.2. Fitting ARIMA models

The ARIMA models are selected using information criterion statistics with an idea of the order borrowed from the ACF and the PACF plots. The best model chosen is one with the least AIC/BIC or the one with the highest Log likelihood function (LLF). Table 2 presents the models with corresponding information criterion statistics.

**Table 2** Fitting ARIMA Models

| **Model** | **AE. Dirham** | | | **Australian Dollar** | | | **Canadian Dollar** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **AIC** | **BIC** | **LLF** | **AIC** | **BIC** | **LLF** | **AIC** | **BIC** | **LLF** |
| ARIMA (0,0,0) | -20343 | -20331 | 10173 | -16198 | -16186 | 8101 | -16692 | -16680 | 8348 |
| ARIMA (0,0,1) | -20354 | -20337 | 10180 | -16210 | -16192 | 8108 | -16690 | -16672 | 8348 |
| ARIMA (1,0,0) | -20352 | -20334 | 10179 | -16208 | -16191 | 8107 | -16690 | -16672 | 8348 |
| ARIMA (1,0,1) | -20358 | -20335 | 10183 | -16227 | -16204 | 8118 | -16725 | -16702 | 8367 |
| ARIMA (1,0,2) | -20367 | -20337 | 10188 | -16226 | -16196 | 8118 | -16734 | -16705 | 8372 |
| ARIMA (1,0,3) | -20384 | -20349 | 10198 | -16224 | -16189 | 8118 | -16743 | -16708 | 8378 |
| ARIMA (2,0,0) | -20371 | -20348 | 10190 | -16214 | -16190 | 8111 | -16695 | -16672 | 8352 |
| ARIMA (0,0,2) | -20367 | -20344 | 10188 | -16217 | -16193 | 8112 | -16697 | -16674 | 8353 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ARIMA (2,0,1) | -20369 | -20340 | 10190 | -16226 | -16196 | 8118 | -16736 | -16706 | 8373 |
| ARIMA (2,0,2) | -20380 | -20345 | 10196 | -16224 | -16188 | 8118 | -16738 | -16703 | 8375 |

Table 2 shows that the KShs/AE. Dirham currency exchange rate return is best modelled by an ARIMA (1,0,3), the KShs/Australian Dollar is fitted best by the ARIMA (1,0,1) model and ARIMA (1,0,3) best fits the KShs/Canadian dollar currency exchange. The fitted models are then tested for adequacy. The Ljung box test was applied and the following results were obtained as shown in Table 3 below.

**Table 3** Ljung-Box test results for residuals

| | AE. Dirham | Australian Dollar | Canadian Dollar |
|---|---|---|---|
| Test Statistic | 17.796 | 10.926 | 7.653 |
| p-value | 0.0554 | 0.3633 | 0.6627 |

The results show that the null hypothesis of the absence of serial correlation or mean effects is not rejected implying that the models are adequate, that is, there exist no remnant mean effects. Moreover, these models can be used to forecast future currency returns for the exchange rates.

Table 4 shows the forecasts for ten days ahead of the Kenyan shilling against the three currencies. The table consists of the forecast and the 95% confidence band for KShs/AE. Dirham, KShs/Australian Dollar and the KShs/Canadian Dollar.

**Table 4** Forecasts of the best fit ARIMA models

| AE. Dirham | | | Australian Dollar | | | Canadian Dollar | | |
|---|---|---|---|---|---|---|---|---|
| Lower Limit | Point Forecast | Upper Limit | Lower Limit | Point Forecast | Upper Limit | Lower Limit | Point Forecast | Upper Limit |
| -0.0099 | 2.4e-4 | 0.0104 | -0.02224 | -1.7e-5 | 0.02221 | -0.02036 | -2.03e-4 | 0.01996 |
| -0.0102 | 2.2e-5 | 0.0102 | -0.02230 | 2.6e-6 | 0.02231 | -0.02005 | 1.2e-4 | 0.02029 |
| -0.0100 | 1.7e-4 | 0.0104 | -0.02232 | 1.6e-5 | 0.02235 | -0.01989 | 3.1e-4 | 0.02052 |
| -0.0102 | 1.5e-5 | 0.0102 | -0.02233 | 2.5e-5 | 0.02238 | -0.02019 | 1.6e-4 | 0.02051 |
| -0.0101 | 1.5e-4 | 0.0104 | -0.02233 | 3.1e-5 | 0.02239 | -0.02028 | 9.4e-5 | 0.02047 |
| -0.0102 | 3.6e-5 | 0.0103 | -0.02233 | 3.6e-5 | 0.02240 | -0.02032 | 6.4e-5 | 0.02045 |
| -0.0101 | 1.3e-4 | 0.0104 | -0.02233 | 3.9e-5 | 0.02240 | -0.02033 | 4.99e-5 | 0.02043 |
| -0.0102 | 5.2e-5 | 0.0103 | -0.02233 | 4.1e-5 | 0.02241 | -0.02034 | 4.4e-5 | 0.02043 |
| -0.0101 | 1.2e-4 | 0.0104 | -0.02233 | 4.2e-5 | 0.02241 | -0.02034 | 4.1e-5 | 0.02042 |
| -0.0102 | 6.2e-5 | 0.0103 | -0.2232 | 4.3e-5 | 0.02241 | -0.02034 | 3.97e-5 | 0.02042 |

### 4.3. Fitting EMD-ARIMA model

First, the data is decomposed into intrinsic mode functions (IMFs) and residuals. Figure 5 presents the IMFs and the residuals of the KShs/AE. Dirham, IMFs and residuals for the KShs/Australian dollar and KShs/Canadian dollar are displayed in Figure 6 and Figure 7 respectively.

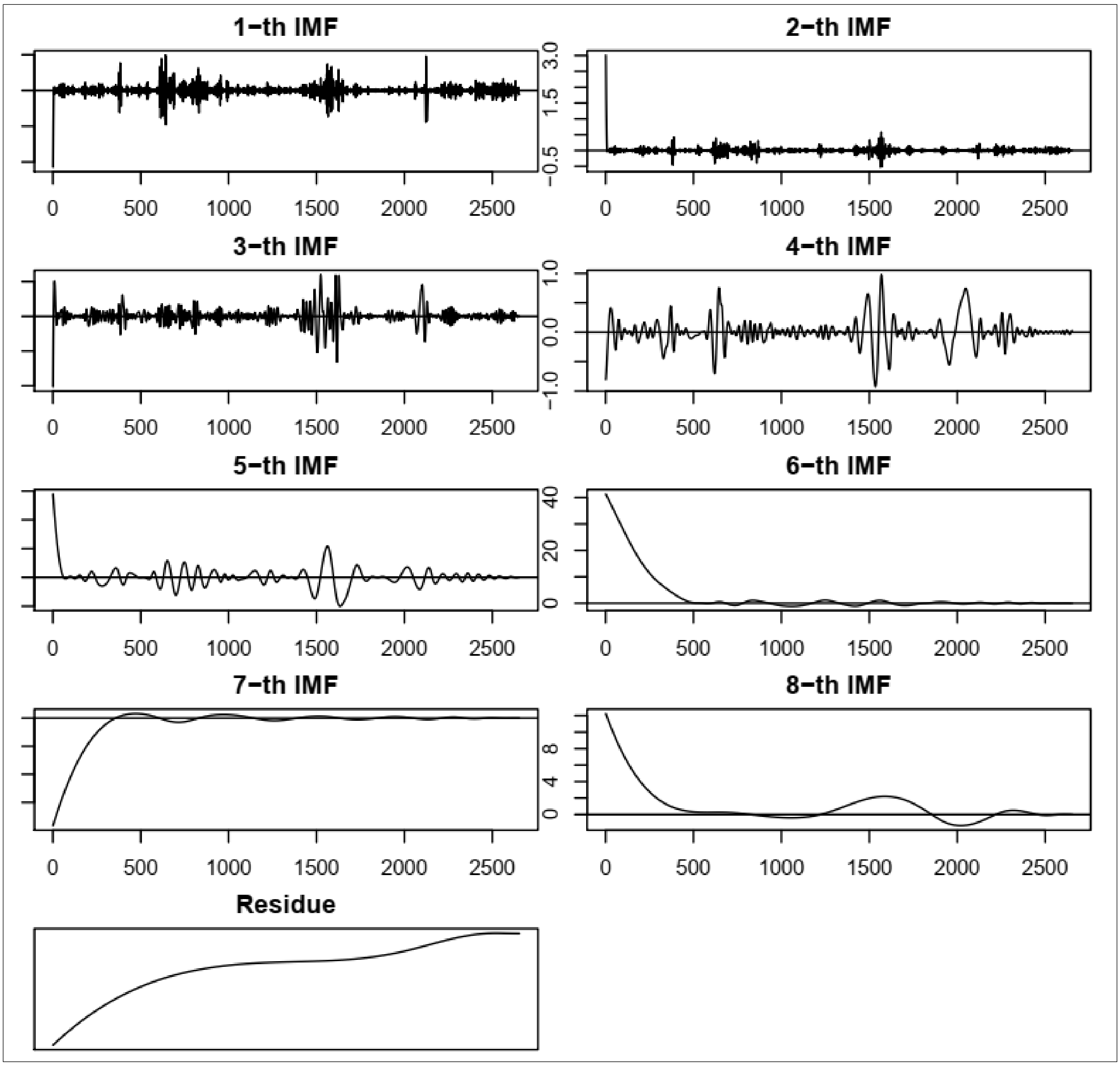


**Figure 5** IMFs and Residue plots for the KShs/AE. Dirham

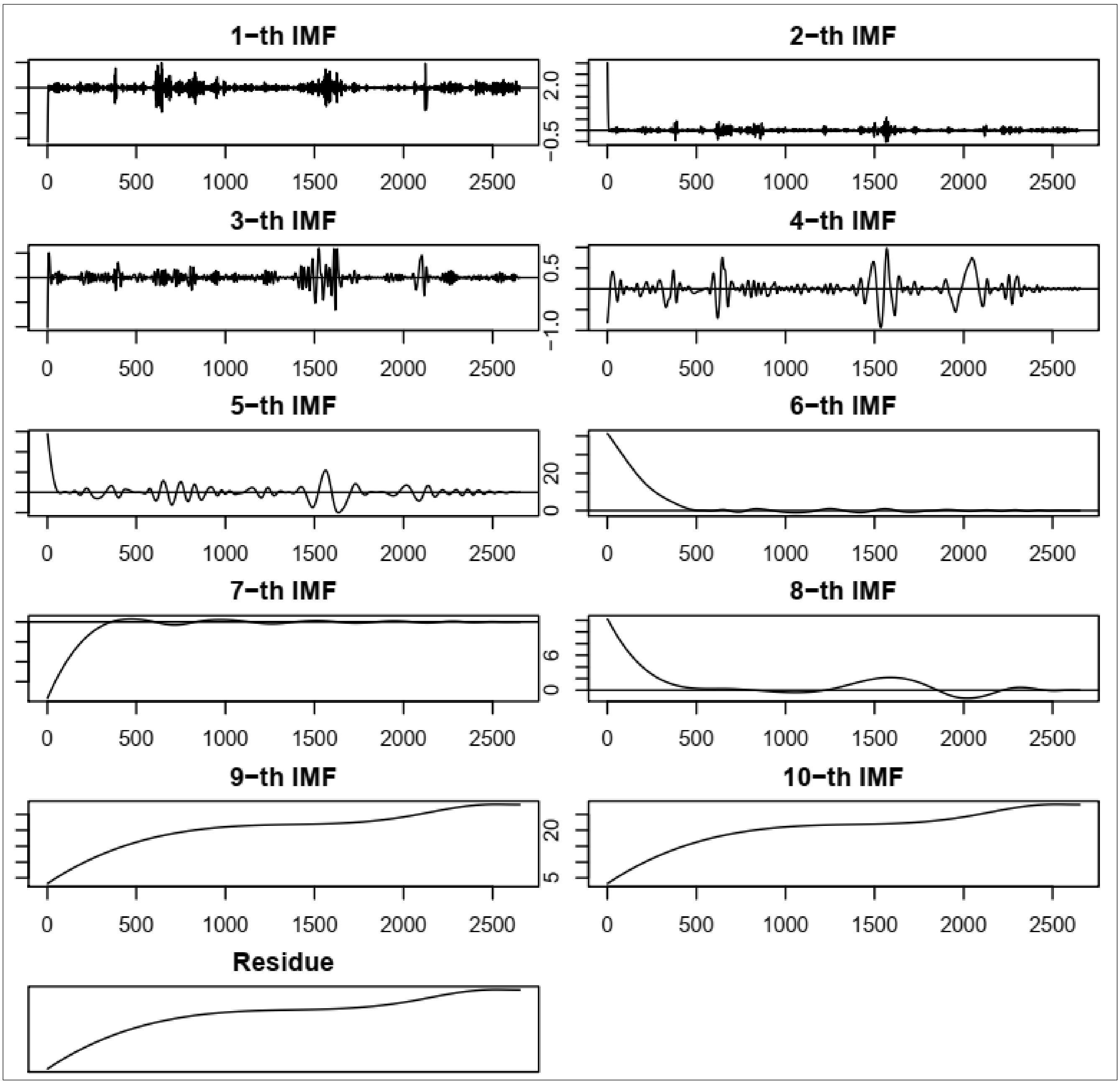


**Figure 6** IMFs and Residue plots for the KShs/Australian dollar

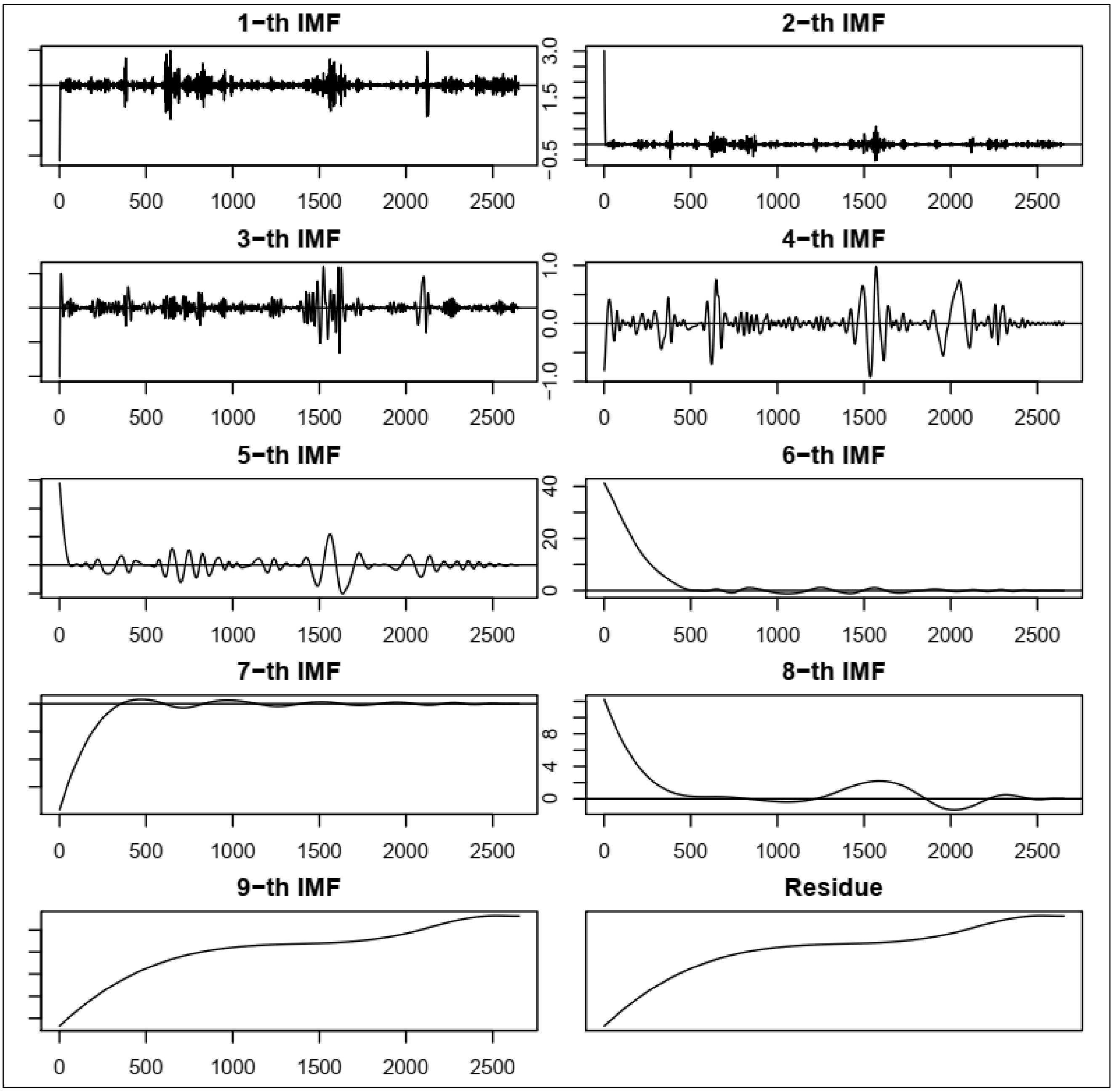


**Figure 7** IMFs and Residue plots for the KShs/Canadian dollar

### 4.4. Comparing the accuracy

Further on, ARIMA models are fitted based on the empirically decomposed data. The models are then compared, the best model being the one with the least AIC/BIC or the highest LLF. Table 5 displays the EMD-ARIMA models with their constituent information criterion statistics for the three currency rates. Furthermore, the models are fitted with the one with the least AIC/BIC or the highest LLF is chosen as the best model. The best EMD-ARIMA for KShs/AE. Dirham, KShs/Australian dollar and KShs/Canadian dollar are EMD-ARIMA (1,0,1), EMD-ARIMA (0,0,1) and EMD-ARIMA (1,0,1) respectively. Since they are the best-fitted models, the next step involves forecasting for several days ahead, in this case, ten days ahead, as forecasting further, the prediction tends to be less reliable.

**Table 5** Fitted ARIMA models

| Model | AIC | BIC | LLF |
|---|---|---|---|
| AE. Dirham | | | |
| ARIMA (1,0,1) | -21652 | -21635 | 10829 |
| ARIMA (0,2,0) | -81903 | -81897 | 40953 |
| ARIMA (2,2,2) | -62550 | -62521 | 31280 |
| ARIMA (0,2,5) | -89969 | -89934 | 44991 |
| Australian Dollar | | | |
| ARIMA (0,0,0) | -25438 | -25432 | 12720 |
| ARIMA (0,0,1) | -17743 | -17731 | 8873 |
| ARIMA (0,2,0) | -77037 | -77031 | 38520 |
| ARIMA (5,0,4) | -23305 | -23246 | 11662 |
| Canadian Dollar | | | |
| ARIMA (1,0,1) | -17938 | -17920 | 8972 |
| ARIMA (0,1,0) | -47645 | -47634 | 23825 |
| ARIMA (0,2,0) | -73042 | -73036 | 36522 |
| ARIMA (0,2,5) | -82352 | -82317 | 41182 |

**Table 6** EMD-ARIMA forecasts

| Currency | AE. Dirham | Australian Dollar | Canadian Dollar |
|---|---|---|---|
| **n-ahead** | **ARIMA (1,0,1)** | **ARIMA (0,0,1)** | **ARIMA (1,0,1)** |
| 1 | 0.004506 | -0.01753 | -0.00966 |
| 2 | 0.005399 | -0.01379 | -0.016784 |
| 3 | 0.01165 | -0.00268 | -0.008116 |
| 4 | -0.002264 | -0.00055 | -0.001328 |
| 5 | -0.002787 | 0.00023 | 0.001615 |
| 6 | -0.001578 | 0.00322 | 0.002534 |
| 7 | -0.00577 | 0.00571 | 0.002461 |
| 8 | -0.000633 | 0.00603 | 0.002038 |
| 9 | -0.001316 | 0.00547 | 0.001635 |
| 10 | -0.001779 | 0.00525 | 0.001400 |

**Table 7** Root mean squared errors of the two models

| Model | AE. Dirham | Australian Dollar | Canadian Dollar |
|---|---|---|---|
| ARIMA | 0.005162 | 0.01138 | 0.010319 |
| EMD-ARIMA | 0.00113612 | 0.01136 | 0.0100211 |

The comparison between the ARIMA model and the hybrid model of EMD-ARIMA is made to ensure the capability of the proposed model to perform better in forecasting exchange rate data. Table 7 shows that the RMSE for each currency in the EMD-ARIMA model is smaller than that of the ARIMA model. It is clearly shown that EMD-ARIMA produces better results than the ARIMA model where it gives the smallest error value in terms of RMSE. Based on this result, there is no doubt that the decomposition technique by using EMD is efficient in improving the performance of the forecasting model.

## 5. Conclusion

Financial time series forecasting plays an important role in the world's economy due to its ability to forecast economic benefits and influence countries' economic development. Currency exchange rates are of interest to financiers and investors. The constant movement of currency exchange rates combined with the rapid globalization of business has resulted in the demand for forecasting methods. However, it is known that financial time series forecasting has shortcomings such as non-linearity and non-stationarity. Therefore, financial time series forecasting is a challenging task in financial markets.

In this paper, the Autoregressive Integrated Moving Average model was used to forecast currency exchange rates. Raw data was first transformed using Empirical Mode Decomposition then a forecast of the exchange rates was done using the ARIMA model. EMD is suitable for financial time series data in terms of finding fluctuation tendency, which then simplifies the forecasting task into several forecasting sub-tasks. EMD as a data processing technique offers a way by which the non-stationary and non-linear characteristics of raw currency exchange rate data can be decomposed into several Intrinsic Mode Functions.

The dataset consists of three foreign exchange rates, that is, KShs/AE. Dirham, KShs/Australian Dollar and KShs/Canadian Dollar from January 2005 to May 2017 daily observations excluding weekends and public holidays. The best fit for the KShs/AE. Dirham, KShs/Australian Dollar and KShs/Canadian Dollar are ARIMA (1,0,3), ARIMA (1,0,1) and ARIMA (1,0,3) respectively. For the EMD-ARIMA model, the best fit for the KShs/AE. Dirham, KShs/Australian Dollar and KShs/Canadian Dollar are ARIMA (1, 0,1), ARIMA (0,0,1) and ARIMA (1,0,1) respectively.

There has been increasing attention given to finding an effective model to address the problem of financial time series forecasting in terms of non-linear and non-stationary characteristics. In this paper, an EMD-ARIMA forecasting model is proposed. EMD is used to detect the moving trend of financial time series data and improve the forecasting success of ARIMA. The EMD-ARIMA model results show that with the implementation of decomposition strategy via EMD to the exchange rate data, the non-linear and non-stationary behavior of the exchange rate data can be addressed effectively and the hidden pattern of the data can be revealed for better understanding resulting in improving the forecasting accuracy. This can be proved with better forecasting results produced by EMD-ARIMA compared to ARIMA. Thus, it can be concluded that the proposed EMD-ARIMA model may be an effective tool for financial time series forecasting.

## Compliance with ethical standards

### *Disclosure of conflict of interest*

No conflict of interest is to be disclosed.

## References

[1] Appiah, S T and I A Adetunde (2011). "Forecasting Exchange Rate Between the Ghana Cedi and the US Dollar using Time Series Analysis". In: 3.2, pp. 76–83.

[2] Hong, Lei (2011). "Decomposition and Forecast for Financial Time Series with High-frequency Based on Empirical Mode Decomposition". In: Energy Procedia 5, pp. 1333–1340. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2011.03.231. URL: http://dx.doi.org/10.1016/j.egypro.2011.03. 231.

[3] Izzati, Nur, Abdul Rashid, Ruhaidah Samsudin, Ani Shabri, and Fakulti Komputeran (2016). "Empirical Mode Decomposition and Least Squares Support Vector Machine". In: 8.3.

[4] Lin, Chiun-sin, Sheng-hsiung Chiu, and Tzu-yu Lin (2012). "Empirical mode decomposition – based least squares support vector regression for foreign exchange rate forecasting". In: Economic Mod- elling 29.6, pp. 2583–2590. ISSN: 0264-9993. DOI: 10.1016/j.econmod.2012.07.018. URL: http://dx.doi.org/10.1016/j.econmod.2012.07.018.

[5] Mong, Tran and Uyen Ngan (2016). “Forecasting Foreign Exchange Rate by using ARIMA Model : A Case of VND / USD Exchange Rate”. In: 7.12, pp. 38–44.

[6] Nwankwo, Steve C (2014). “Autoregressive Integrated Moving Average ( ARIMA ) Model for Exchange Rate ( Naira to Dollar )”. In: 3.4, pp. 429–434. DOI: 10.5901/ajis.2014.v3n4p429.

[7] Yu, Lean, Shouyang Wang, and Kin Keung (2008). “Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm”. In: 30, pp. 2623–2635. DOI: 10.1016/j.eneco. 2008.05.003.

[8] Zeng, Q., & Qu, C. (2014). An approach for Baltic Dry Index analysis based on empirical mode decomposition. Maritime Policy & Management, 41(3), 224-240.